Discover Artificial Intelligence

REVIEW Open Access

Zero-click AI epistemic injustice and the governance of digital knowledge infrastructures

Selcen Ozturkcan[1,2*]

*Correspondence:
Selcen Ozturkcan
selcen.ozturkcan@lnu.se; selcen@sabanciuniv.edu
[1]School of Business and Economics, Linnaeus University, Växjö, Sweden
[2]Sabanci Business School, Sabanci University, Istanbul, Turkey

Abstract

This paper examines *zero-click AI,* systems that deliver synthesized answers directly within platform interfaces without prominently visible source-level provenance. Such designs restructure conditions of credibility, visibility, and authorship by displacing upstream creators, suppressing provenance, and concentrating authority. I argue that these systems position platforms as functional epistemic agents whose design choices reliably generate epistemic harms. To analyze this, I develop a three-layer framework, focusing on infrastructures, appropriation, and governance, that specifies the mechanisms by which epistemic injustices arise under prevailing configurations. Building on debates in epistemic injustice theory, I show how zero-click systems systematically efface attribution, marginalize plural knowledges, and entrench epistemologies of ignorance. The normative argument is conditional rather than metaphysical: harms are contingent on design and policy logics, not inevitable features of automation. On this basis, I propose reform thresholds, attribution by default, provenance affordances, redistributive mechanisms, and pluralistic sourcing, derived from principles of proportionality and value-sensitive design. Properly operationalized, these measures demonstrate how digital infrastructures might preserve accessibility and efficiency while sustaining the pluralism required for a resilient knowledge commons.

Keywords Zero-click AI, Epistemic injustice, Knowledge commons, Epistemic agency, Platform enclosure, Philosophy of technology

1 Introduction

Imagine a student who turns to a search engine for information about a health condition. Instead of a list of articles by doctors, medical associations, or science communicators, she sees a concise AI-generated summary inside the interface. The author is absent, the source invisible, and the click unnecessary. I call this pattern zero-click collapse, not to claim an inevitable or total breakdown of the knowledge commons, but as a heuristic for a structural contraction of attribution, reciprocity, and pluralism under current platform logics. By *collapse,* I therefore mean not systemic breakdown but a patterned narrowing of visible authorship, outward navigation, and epistemic traceability within

interface defaults that centralize authority in the platform. The harms are conditional, not fated: when provenance is suppressed and attention is enclosed, human voices are overwritten by synthetic synthesis, and the anchors that once tied knowledge to identifiable communities of expertise recede from view.

This phenomenon, which I describe as *zero-click collapse*, cannot be reduced to interface design. It signals a structural transformation of digital infrastructures in which platforms no longer merely index and distribute knowledge but actively recompose and reallocate epistemic authority. Where the hyperlink once sustained transparency and pluralism, the zero-click model centralizes attention, marginalizes content creators, and positions the platform itself as an epistemic agent. Such a reconfiguration of informational flows raises urgent philosophical questions about knowledge, power, and justice in the digital age.

While public debate around generative AI often emphasizes novelty, zero-click collapse is better understood as a continuation and intensification of the longer trajectory of platform capitalism. The economic logic of enclosure, capturing attention and reducing outward flows, has long shaped digital ecosystems [1], [2]. With the adoption of generative AI, these dynamics deepen: large models ingest human-authored content and return it in platform-native form, frequently without attribution or reciprocity [3], [4]. The result is not merely economic disruption to creative and educational labor but also an epistemic reorganization in which platforms arbitrate visibility, credibility, and trust.

Philosophical reflection is indispensable in confronting these transformations. Zero-click systems exemplify how technological infrastructures are never neutral but embody normative commitments that shape not only what is known but also how it comes to be known. By compressing diverse perspectives into seemingly authoritative outputs, such systems risk epistemic injustice [5], narrowing the horizons of public discourse and effacing the social labor on which AI depends. This makes zero-click collapse paradigmatic for the philosophy of technology: it is a site where design decisions restructure epistemic relations, where ethical dilemmas of attribution and appropriation emerge, and where the political economy of knowledge is recalibrated under conditions of enclosure.

Critical AI ethics has rightly stressed structural harms, opacity, and power asymmetries [6],[7], [8], [9]. While these accounts foreground the political economy and ethical regulation of AI, they have not theorized the epistemic injustices that arise under specific design and governance conditions. The framework here complements those critiques by specifying epistemic mechanisms and justice-oriented thresholds.

The central contribution of this paper is to conceptualize zero-click collapse as the manifestation of platforms as epistemic agents: socio-technical systems that mediate and redistribute epistemic authority [10], [11]. In using this term, I adopt a functional–relational notion of agency (after [12, 13]): platforms exercise epistemic agency insofar as their design choices and defaults systematically shape uptake, credibility, and trust. Responsibility remains attributable to the human and organizational actors who specify, deploy, and govern these systems [11]. To develop this argument, I propose a three-layer analytic framework, namely platform design, knowledge commons, and governance, that structures a philosophical analysis of how digital infrastructures reconfigure economic incentives, ethical relations, and epistemic conditions [14]. The claim advanced is conditional, not metaphysical: epistemic injustices arise reliably when three sufficient

conditions co-obtain: (i) answer synthesis lacks prominently visible source-level provenance, (ii) generative substitution displaces upstream authors in user attention, and (iii) ranking and revenue policies preclude reciprocity and contestation. Importantly, the normative demand for attribution advanced here does not depend on material sameness or copyright-level copying but on how platforms present synthesized outputs as authoritative while suppressing visible epistemic lineage. By situating zero-click collapse within debates on technological mediation, epistemic injustice, and AI ethics, the article extends the philosophy of technology's concern with the normative dimensions of design into the terrain of digital knowledge infrastructures. The paper clarifies when platform couplings count as functional epistemic agency (without anthropomorphism), integrates epistemic injustice into infrastructural design by identifying conditions under which such injustices are systematically likely, and advances normative thresholds as philosophical criteria for accepting or rejecting zero-click designs.

By epistemic agency I mean not literal authorship or consciousness, but functional agency in Verbeek's [13], [15] and Sundar's [12] sense: platforms configure conditions of uptake, credibility, and trust such that they are treated as authoritative speakers. Responsibility remains attributable to the human and organizational actors who design and govern these systems, yet functionally the platform itself is positioned as the locus of epistemic authority.

The next section develops the three-layer framework of epistemic infrastructure, appropriation, and governance, and specifies what counts as zero-click AI. The following analysis examines platform economies and epistemic authority, ethical opacity and epistemic injustice, governance asymmetries, reform thresholds, and proportionality in relation to countervailing goods. The conclusion synthesizes the implications for technological mediation, epistemic agency, and the future of digital knowledge.

2 The architecture of collapse: epistemic infrastructures, appropriation, and governance in zero-click systems

For the purposes of this article, zero-click AI refers to generative AI-mediated information delivery that (i) presents synthesized answers directly within the platform interface, (ii) defaults to in-situ consumption rather than outward navigation, and (iii) provides no prominently visible, source-level provenance that would enable users to trace contributions. This formulation distinguishes zero-click AI from earlier snippet functions, which quoted sources verbatim, and from enterprise question–answer systems, which operate within bounded organizational contexts and contractual regimes. The concern here is with public-facing infrastructures that mediate general-purpose knowledge access at scale. Where these design and policy conditions are absent or mitigated, the risks of epistemic injustice discussed below do not necessarily obtain, underscoring the conditional rather than inevitable character of the harms. Zero-click AI excludes (i) snippet cards that present verbatim quotations with prominently visible, above-the-fold citations and link-outs to the first-party source and (ii) purely navigational tiles that do not synthesize third-party content.

To grasp the systemic consequences of zero-click architectures, it is necessary to move beyond critiques of user interface design or economic disruption and toward a conceptual account of how these systems mediate knowledge itself. Platforms are not passive conduits but infrastructures that actively redistribute epistemic authority, shaping what

can be known, how it is valued, and which voices are amplified or erased [16], [13], [17]. Zero-click collapse, therefore, requires an analytic framework that captures its economic, ethical, and epistemic dimensions as forms of technological mediation.

The three-layer framework advanced here provides such a perspective.

- The first layer is that of epistemic infrastructures, where platform design choices, answer synthesis, attention capture, suppression of provenance, and enclosed attention and position the platform as an epistemic agent [10, 12].
- The second is the knowledge commons, where the appropriation of human-authored content for AI training can generate epistemic asymmetries, displace creative labor, and reduce plural contributions [5, 18].
- The third is governance, where institutional arrangements, often opaque and centralized, may conflict with principles of inclusivity, transparency, and accountability associated with value-sensitive and participatory governance [19], [14].

These relationships are represented in Fig. 1, which visualizes the architecture of zero-click collapse across the three interconnected layers. At the level of infrastructures, platforms enclose attention through answer synthesis and provenance suppression. At the level of appropriation, the commons is drawn upon without reciprocity, leading to asymmetry and the effacement of authorship. At the level of governance, opaque and centralized mechanisms may conflict with democratic, value-sensitive, and participatory governance principles. The figure does not depict an empirical model but a conceptual lens for interrogating the epistemic dynamics of zero-click systems. The term "*layers*" is used here analytically rather than ontologically: these dimensions are not discrete strata but interdependent aspects of the same socio-technical configuration. The triangular visualization is heuristic, intended to represent structural interconnection rather than hierarchy or mutual exclusivity.

I use epistemic agency in a two-level sense.

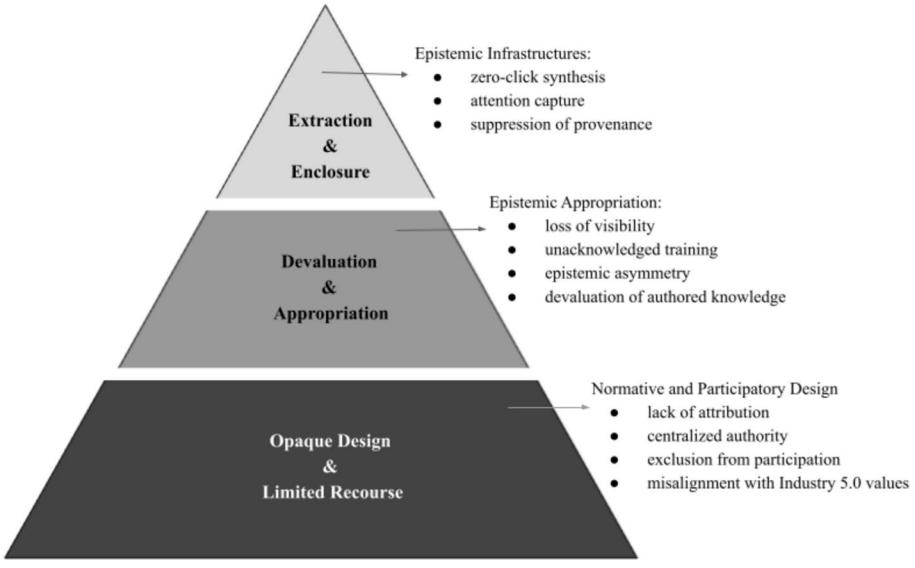

Fig. 1 Conceptual architecture of zero-click collapse: three layers of epistemic infrastructures, appropriation, and governance. (Author's own illustration)

1. Functional/relational agency. Following Verbeek's account of technological mediation and Sundar's notion of machine agency, platforms, understood as assemblages of interfaces, ranking/retrieval models, policies, and operator practices, co-produce epistemic outcomes by (i) curating and synthesizing content, (ii) configuring defaults that channel attention, and (iii) assigning credibility signals that guide user uptake. This is a claim about the systemic role in shaping justification and trust, not about autonomy or mental states [12, 13].
2. Attributable agency. Responsibility for the epistemic effects of these systems is ascribed to human and organizational actors, designers, product owners, executives, and regulators, who specify objectives, training data, ranking/logging rules, and revenue policies [11], [10].

Extending agency talk from artifacts to platforms is justified because platforms are stabilized socio-technical assemblages whose recurrent couplings (UI patterns, recommender parameters, moderation and monetization rules) produce predictable epistemic regularities. This functional use avoids equivocation: when the paper says platforms act as epistemic agents, it means they play an action-guiding role in the distribution of credibility and uptake; when it assigns obligations, it refers to attributable agency borne by platform controllers and governance arrangements.

To avoid overstating the claim, it is important to stress that epistemic injustices are not metaphysically necessary consequences of zero-click AI. These conditions arise reliably when specific design and governance factors are present simultaneously. Table 1 aligns mechanisms across the three layers with their triggering conditions, the specific epistemic harms they produce, and the normative implications that follow. This analytic mapping anchors the argument in concrete pathways, distinguishing necessary from contingent harms and clarifying how structural features of zero-click systems systematically tend toward epistemic injustice. By mechanisms, I refer to recurrent socio-technical configurations, e.g., combinations of design defaults, ranking logics, and governance policies, that reliably generate patterned epistemic effects under specified conditions.

2.1 Epistemic infrastructures and platform mediation

Zero-click content delivery represents not merely a change in user interface but a reconfiguration of epistemic infrastructures. Historically, the hyperlink anchored web

Table 1 Mechanisms of zero-click collapse: conditions, epistemic harms, and normative implications

Mechanism (Layer)	Triggering condition	Epistemic harm	Normative upshot
Answer synthesis without visible provenance (Infrastructures)	Sources hidden or secondary	Testimonial injustice; degraded source traceability	Attribution by default; surface-level citations
Generative substitution of upstream content (Appropriation)	Model outputs compete with originals without reciprocity	Hermeneutical injustice; creator displacement	Redistribution mechanisms; provenance affordances
Opaque ranking and revenue policies (Governance)	Closed curation and monetization	Concentration of epistemic authority; marginalization of minority voices	Pluralistic sourcing; public audit and accountability
Data colonialism in training corpora (Appropriation/Governance)	Dominant-language bias; weak inclusion	Epistemic homogenization	Diversity metrics; periodic bias audits
Machine agency shaping user reliance (Infrastructures)	Default single best answer	Over-reliance; reduced critical engagement	Choice architecture with multi-source views

architectures by enabling users to trace claims back to identifiable sources, thereby sustaining pluralism and discoverability [20], [4]. The hyperlink functioned both technically and epistemically as a distributed mechanism of accountability [21], [22]. By contrast, zero-click systems embed synthesized answers within the platform interface, positioning the platform itself as the immediate locus of epistemic authority [2], [23].

As Star and Ruhleder [24] classically argue, infrastructures become visible upon breakdown, and Edwards [25] shows how opacity is constitutive of global information infrastructures. Zero-click architectures exemplify this dynamic: Provenance suppression and answer synthesis render the infrastructural mediation of knowledge invisible until epistemic harms, displacement of creators, erosion of pluralism, and surface as breakdowns.

On the present account, this constitutes functional epistemic agency. Through defaults, synthesis modules, and ranking mechanisms, platforms co-produce epistemic outcomes by directing attention and structuring credibility, even without autonomous intention [12, 13, 26]. As Ihde [16] and Verbeek [15] have argued, technologies mediate human–world relations, shaping how knowledge is accessed, trusted, and evaluated. Zero-click architectures exemplify this: mediation occurs through enclosure, whereby the platform simultaneously delivers and delimits the field of knowledge.

Extending agency talk to platforms is justified insofar as they are stabilized socio-technical assemblages whose recurrent couplings, interfaces, retrieval models, moderation protocols, and monetization rules generate predictable epistemic regularities [27], [17]. These patterns of mediation allow platforms to function as epistemic agents in a relational sense: they shape justification practices and uptake, while responsibility for their design and operation remains with human and organizational actors [11], [10].

This infrastructural transformation also reflects a broader logic of media convergence, whereby indexing, hosting, synthesis, ranking, and monetization are consolidated within a single platform architecture. Rather than merely aggregating content, AI-driven systems orchestrate machine-to-machine processes, e.g., crawling, training ingestion, retrieval, and generative synthesis, that precede and structure human interaction. In this sense, authority is not only redistributed at the level of interface presentation but also preconfigured within convergent socio-technical assemblages. Read through Ihde's account of technological mediation [16] and extended by postphenomenological analyses of technological agency [12, 13], such convergence reconfigures the conditions under which credibility, attribution, and uptake are structured.

2.2 Epistemic appropriation and the knowledge commons

Zero-click architectures depend not only on infrastructural mediation but also on systematic appropriation of human-authored content. Large training corpora and generative substitution draw on texts, images, and code produced by upstream contributors, often without mechanisms of attribution or reciprocity [3, 28]. This is consistent with broader critiques of data colonialism, in which human life and knowledge are appropriated as raw material for extraction [17], [29].

The result is a knowledge commons in which contributions circulate in derivative form while their origins are obscured. Such practices produce epistemic asymmetries: platforms aggregate and recombine content at scale, while individual creators lose visibility in the chain of recognition [30], [1]. Appropriation here is not incidental but structural,

as training data, synthesis outputs, and retrieval models continuously incorporate human contributions into platform-level outputs [18].

Analytically, this layer highlights how the commons is reconfigured by generative AI. Contributions that once sustained pluralism through link-based visibility are subsumed into aggregated answers, with credit and traceability often severed. This resonates with Fricker's [5] concept of epistemic injustice: testimonial authority is undermined when sources are effaced, and hermeneutical resources are eroded when contributions are subsumed without recognition. Related analyses of misinformation ecosystems show that epistemic harms frequently arise not only from falsehoods but from structural distortions in credibility allocation and authority presentation, where visibility and uptake are redistributed independent of source integrity [31]. Appropriation thus operates as a mechanism by which platforms consolidate epistemic authority: they mediate not only access but also authorship and ownership within digital knowledge infrastructures.

2.3 Normative and participatory design in governance

The third analytic layer concerns the governance of zero-click systems. Platforms do not only mediate knowledge through infrastructures and appropriation; they also establish rules of access, monetization, and oversight that shape epistemic environments. Governance here refers to the policies, contractual arrangements, and technical protocols through which platforms organize participation [27, 32].

These structures are typically centralized and opaque. Decisions about ranking, advertising integration, moderation, and citation display are taken within proprietary frameworks, with limited opportunities for public contestation or audit [1, 33]. As Gillespie [27] notes, governance in platform contexts often operates less as transparent rule-making and more as infrastructural sovereignty, the capacity to unilaterally determine how information circulates.

Analytically, this layer highlights how governance determines which contributions are visible, how revenues circulate, and what forms of accountability are recognized [34], [19]. It provides the institutional scaffolding within which epistemic infrastructures and knowledge commons are configured, and through which platform-level epistemic agency is enacted. Governance is thus the meta-layer through which infrastructural and appropriation dynamics are consolidated into durable socio-technical arrangements [14].

The three-layer framework advanced here builds on, but also extends, existing approaches in philosophy of technology and socio-technical systems (STS). Feenberg's [35] critical theory emphasizes the social shaping of technical codes and the possibility of democratic rationalization. While this article shares his concern with normative contestation, it focuses more narrowly on epistemic harms and testimonial, hermeneutical, and contributory injustices that Feenberg leaves under-specified. Jasanoff's [36] co-productionist framework highlights how knowledge and social order are mutually constituted, yet it is not designed to track the mechanisms of epistemic injustice that emerge in AI-mediated infrastructures. Couldry and Mejias's [17] concept of data colonialism identifies the extractive logics underlying digital economies but does not map the specific conditions under which epistemic authority is redistributed and pluralism eroded in zero-click architectures. The present framework contributes by specifying a layered analytic lens, epistemic infrastructures, appropriation, and governance that distinguishes

necessary from contingent harms, aligns mechanisms with normative implications, and operationalizes reform thresholds. In this way, it reframes zero-click systems not simply as instances of enclosure or extraction but as systematic configurations of epistemic injustice.

Contemporary debates on generative AI have largely centered on intellectual property questions: infringement, derivation, originality thresholds, and economic compensation [37], [38], [39]. These analyses ask whether AI outputs reproduce protectable expression, whether derivative use violates copyright, and whether market substitution justifies redistributive licensing regimes. The normative core of this debate is rights-based and jurisdictionally structured, concerned with authorship, ownership, and economic harm.

The framework advanced here operates on different conceptual terrain. Epistemic injustice and epistemic governance approaches are not primarily concerned with copying or legal protectability, but with authority presentation, credibility allocation, and the visibility of epistemic lineage [40], [41]. The question is not whether an output is legally infringing, but whether platform architectures reconfigure who appears as an authoritative knower and whether upstream contributors retain epistemic standing within the interface.

Thus, even where AI-generated outputs satisfy originality thresholds or qualify as transformed under copyright doctrine, epistemic harms may still arise if platforms suppress provenance and present synthesized answers as self-standing authorities. Legal permissibility does not exhaust normative permissibility. The epistemic injustice frame captures governance-level harms that remain conceptually invisible within infringement-based analysis.

To illustrate, consider a time-sensitive health query in which a platform presents an in-interface synthesis that draws on multiple public sources. Assume, arguendo, that the output is legally permissible, that is non-infringing, sufficiently transformed, or licensed, so that copyright doctrine would treat it as original or otherwise allowable. Even under that assumption, the epistemic injustice diagnosis still activates: the interface presents the synthesis as a self-standing authority while relegating source-level provenance to non-salient locations (or omitting it), thereby reallocating credibility from identifiable expert communities to the platform and depriving users of the interpretive resources needed to evaluate standing and contestation. What becomes visible here is not a rights violation but a patterned suppression of epistemic lineage and a redistribution of epistemic credit and trust. The analytic yield is therefore governance-level: it identifies harms to credibility allocation and pluralism that can persist even when infringement-based analysis finds no fault.

3 Discussion: philosophical stakes of zero-click collapse

The analytic mapping above clarifies how infrastructural design, appropriation practices, and governance arrangements interact to produce distinctive epistemic dynamics. The next step is to examine the normative significance of these findings. The same mechanisms that enclose attention, displace creators, and concentrate authority also restructure conditions of visibility, credibility, and participation in the digital public sphere. In this sense, zero-click systems implicate not only economic and informational ordering but also broader questions of epistemic justice, democratic pluralism, and collective flourishing.

Section 2 distinguished infrastructures, appropriation, and governance as the constitutive layers of zero-click architectures. Each of these layers does more than organize technical operations; they instantiate normative stakes that directly implicate epistemic justice, democratic accountability, and human-centered design. Section 3 now turns to these normative consequences. Each subsection addresses a distinct dimension: the redistribution of epistemic credit (3.1), the appropriation of contributions (3.2), the governance asymmetries that structure pluralism (3.3), the operationalization of reforms (3.4), and the proportionality of countervailing goods such as accessibility or efficiency (3.5).

Philosophically, the central concern is that zero-click systems displace epistemic credit and reconfigure the conditions under which knowledge is trusted and valued. By severing contributions from attribution, subordinating human authors to derivative outputs, and concentrating decision-making power within opaque corporate governance, these systems enact what Fricker [5] terms testimonial and hermeneutical injustice. Testimonial injustice arises when the credibility of sources is unfairly withheld or obscured, while hermeneutical injustice occurs when interpretive resources are structurally denied.

These harms are amplified by what Couldry and Mejias [17] conceptualize as data colonialism: the extraction and appropriation of human contributions as raw material for machine outputs. In this sense, the knowledge commons becomes a field of exploitation, where plural and situated knowledges are subsumed under homogenized outputs optimized for platform retention. Similar critiques are raised by Santos-Gonçalves [30], who highlights how independent knowledge producers are structurally marginalized, and by Safadi and Watson [1], who document how opaque curation entrenches asymmetries of epistemic power.

Beyond testimonial and hermeneutical harms, zero-click architectures also risk contributory injustice [42], where marginalized knowers' interpretive resources are appropriated without uptake, and marginal knower exclusion [43], where standpoint-specific perspectives fail to register within platform logics. A virtue-theoretic lens [44] further shifts part of the normative burden to audience- and designer-side responsibilities. These complementary frameworks, e.g., epistemic injustice [5], data colonialism [17], epistemologies of ignorance [45, 46], and resistance [44], provide the scaffolding for the subsections that follow, which evaluate how contingent but systematic harms emerge across layers and what reforms might realign these infrastructures with epistemic justice.

3.1 Platforms as epistemic agents: economy and authority

At the heart of zero-click architectures lies a reconfiguration of the material and epistemic conditions that have historically sustained independent knowledge production. In the hyperlink economy, creators, whether investigative journalists, educators, or science communicators, could expect visibility and revenue through outward navigation, where attention translated into advertising income, subscriptions, or reputational capital [23], [20]. Zero-click delivery interrupts this circuit by embedding synthesized answers directly in platform interfaces, depriving creators of traffic and, crucially, of the epistemic credit tied to attribution [28],[30]. At the user level, this restructuring also preserves the appearance of informational convenience while narrowing deliberative engagement, a pattern consistent with analyses of *choice without choosing* in algorithmically mediated environments [47].

The normative consequences extend beyond financial loss. Local journalism offers a paradigmatic case: minority-serving and rural newsrooms, already precarious, have their work extracted as input for generative summaries that displace user engagement rather than direct it. Their reporting is made visible as content, but their institutional presence becomes invisible, weakening editorial independence and eroding democratic infrastructures of accountability [1]. Educational and scientific content providers face similar dynamics, where tutorials, open-access materials, or STEM resources are subsumed into synthetic outputs, leaving authors without recognition or reputational standing [4]. In each instance, epistemic credit, the acknowledgement of authorship, expertise, and interpretive labor, is redistributed upward to platforms.

Philosophically, this displacement exemplifies how infrastructures mediate epistemic authority. Ihde [16] and Verbeek [13], [15] have shown that technologies co-constitute human–world relations, shaping what counts as knowledge and how it circulates. Coeckelbergh [10] extends this insight by arguing that technologies can function as epistemic agents insofar as they guide conditions of credibility and uptake. Platforms exercise such functional epistemic agency through synthesis, ranking, and interface defaults, while attributable agency, the responsibility for these effects, remains with designers, executives, and regulators who configure objectives and policies [11].

The result is an epistemic asymmetry: creators lose not only livelihood but also the conditions under which their authority is recognized. Couldry and Mejias [17] frame this as data colonialism, in which human knowledge and life are appropriated as raw resources, while Mackenzie [2] demonstrates how platform logics of enclosure restructure the political economy of attention. In zero-click systems, these dynamics culminate in a dual displacement: platforms appropriate both economic value and epistemic standing, consolidating themselves as arbiters of meaning.

Thus, platforms as epistemic agents are not neutral mediators of convenience but infrastructures that reallocate authorship, credibility, and authority. This realignment underscores the central normative concern: economic disenfranchisement is inseparable from epistemic marginalization, and both are systematically produced under prevailing zero-click design and governance conditions.

3.2 Ethical opacity and epistemic injustice

If infrastructure positions platforms as epistemic agents, appropriation exposes how such agency is exercised through asymmetrical extraction. Zero-click systems systematically depend on the unacknowledged incorporation of human-authored content into generative training and synthetic outputs. What was once a commons structured by hyperlink visibility is transformed into a reservoir of raw material, where creators' contributions circulate without attribution or reciprocity [28].

Alcoff [48] underscores how epistemic authority is socially distributed, a point sharpened here by showing how zero-click systems reallocate authority upward to platforms. Mills's [46] work on epistemologies of ignorance likewise illuminates how design choices can actively produce structural unknowing,users are denied the interpretive resources to situate claims when provenance is suppressed.

The normative stakes lie in opacity. Generative substitution displaces upstream authors in user attention, but the conditions of that displacement remain largely invisible to both creators and publics. Training datasets, synthesis boundaries, and ranking

policies are rarely disclosed, producing a form of epistemic opacity that denies stakeholders the resources needed to contest or even understand appropriation [3]. This asymmetry undermines recognition: creators lose epistemic credit, and users lose the interpretive context necessary to situate knowledge claims.

Philosophically, these dynamics instantiate what Fricker [5] terms "*hermeneutical injustice*," where interpretive resources are structurally denied. Creators cannot make sense of how their contributions are reframed, while audiences cannot evaluate credibility without visible provenance. Testimonial injustice follows in turn, as credibility is unfairly redistributed from human authors to synthetic outputs. The knowledge commons, traditionally a site of plural contribution and mutual recognition, is thus hollowed out into what Couldry and Mejias [17] call data colonialism: a space of extraction in which situated knowledges are subsumed under platform authority.

The consequences are not evenly distributed. Santos-Gonçalves [30] documents how independent journalists and educators are particularly vulnerable to displacement, while Wu, Hu, and Chen [4] show how STEM educators' resources are devalued in derivative outputs. Safadi and Watson [1] note that opaque curation disproportionately marginalizes minority voices and non-Western epistemologies. These harms highlight how appropriation reshapes the commons along lines of exclusion, privileging dominant corpora while effacing diversity.

Normatively, opacity in appropriation practices raises questions of accountability. This opacity is a live instance of contributory injustice: creators' situated insights are mined but not credited or allowed to shape uptake [42, 43]. If creators cannot trace how their labor sustains platform outputs, and users cannot verify provenance, the very conditions of epistemic justice collapse. Minimal safeguards, attribution by default, provenance affordances, and redistributive mechanisms are therefore not optional enhancements but ethical requirements for sustaining a viable knowledge commons [33, 34]. Without them, appropriation ceases to be a practice of knowledge sharing and becomes a mechanism of epistemic enclosure.

3.3 Governance and epistemic asymmetry

If economic disenfranchisement exposes the material vulnerabilities of creators and ethical opacity reveals the erasure of provenance, governance discloses the structural asymmetries that entrench platform authority. Zero-click architectures do not simply follow from technical affordances; they are embedded in regimes of control where decisions about visibility, attribution, and redistribution are monopolized by platform providers. Governance here operates less as transparent rule-making than as infrastructural sovereignty: the unilateral capacity to determine how knowledge circulates, under what terms it is accessed, and whose voices are amplified or effaced [1, 27, 49]. Such patterns of centralized control risk devolving into what has elsewhere been described as cycles of AI washing and AI booing, where symbolic gestures of responsibility substitute for structural redesign, thereby intensifying rather than resolving legitimacy deficits [50].

These asymmetries are not metaphysically inevitable, but they arise reliably under prevailing design logics. Attributable agency rests with platform controllers, who specify ranking policies, monetization rules, and provenance displays. Functional epistemic agency, in Verbeek's [13], [15] and Sundar's [12] sense, is exercised when interfaces, retrieval models, and defaults structure how credibility and authority are distributed.

Table 2 Mapping mechanisms to injustices, reforms, and normative justifications

Mechanism (Layer)	Epistemic injustice produced	Reform proposal	Normative justification
Answer synthesis without visible provenance (Infrastructures)	*Testimonial injustice* (creators' credibility effaced); *Hermeneutical injustice* (users lack interpretive resources) [5]	Attribution by default (inline link-outs to ≥ 3 sources, with visible author/outlet/year)	Restores fair credibility distribution [33], meets minimal sufficiency for trust evaluation [34]
Generative substitution of upstream content (Appropriation)	*Contributory injustice* (marginalized knowers excluded from interpretive resources) [42], economic disenfranchisement	Redistributive mechanisms (transparent allocation key; platform-funded compensation)	Responsibility by design falls on controllers, not end-users [11]; supports epistemic reciprocity
Opaque ranking and governance policies (Governance)	*Hermeneutical injustice* (interpretive opacity); *Epistemologies of ignorance* entrenched [44, 45]	Provenance affordances (expandable panels showing excerpt boundaries, model confidence, type of use)	Provides interpretive sufficiency for publics [34], operationalizes Verbeek's [15] view of design as normative
Data colonialism in training corpora (Appropriation/Governance)	*Epistemic homogenization;* exclusion of minority languages and situated knowledges [17] [46]	Pluralistic sourcing & audit (metrics on language, geography, outlet diversity; periodic public audits)	Implements Medina's [44] epistemic virtues of resistance,aligns with Helberger & Wojcieszak [51] on exposure diversity
Platform defaults as machine agency (Infrastructures)	Over-reliance; reduced critical engagement [12]	Proportionality principle (least-restrictive means: inline citations, expandable provenance, multi-source views)	Ensures accessibility/speed are balanced with justice; prevents convenience from devolving into epistemic enclosure

Harms occur when these levels of agency converge without countervailing safeguards, concentrating epistemic authority and marginalizing plural voices.

The normative significance of this concentration follows directly from the epistemic injustices mapped earlier. First, it undermines epistemic pluralism: minority-language contributions, situated knowledges, and non-Western epistemologies are systematically excluded [17] [46]. Second, it entrenches testimonial injustice by stripping creators of credibility through effaced attribution and hermeneutical injustice by denying publics the interpretive resources to situate claims [5]. Third, it produces contributory injustice [42], where marginalized knowers are structurally prevented from contributing interpretive resources that could enrich shared understanding. These harms are intensified by epistemologies of ignorance [44, 45], which flourish under opaque governance. Crucially, these are contingent harms: where provenance, reciprocity, and pluralistic governance are embedded, they do not arise.

3.4 From justice analysis to reform

The following reform proposals are not presented as freestanding policy prescriptions but as normative corollaries of the preceding justice analysis. Each measure corresponds to a specific mechanism of epistemic harm identified in Table 2 and is justified as the least-restrictive design intervention capable of addressing testimonial and hermeneutical

injustices without negating countervailing goods. In this sense, attribution, provenance, redistribution, and pluralistic sourcing are derived from philosophical argument rather than appended as policy advocacy.

- *Attribution by default responds to testimonial injustice.* Without visible credit, authors' credibility is structurally suppressed. A minimally sufficient criterion is inline link-outs to at least three distinct sources per generated answer, with visible cues (author, outlet, year). This implements Brand's [33] attribution by design principle and restores the conditions for fair credibility distribution.
- *Provenance affordances address hermeneutical injustice.* Users require interpretive resources to situate claims; token acknowledgments are insufficient. Expandable panels disclosing excerpt boundaries, model confidence, and use type (quotation/paraphrase/training) satisfy Joyce et al.'s [34] demand for interpretive sufficiency, giving publics the tools to understand epistemic lineage.
- *Redistributive mechanisms target economic disenfranchisement and contributory injustice.* By tying compensation to actual use through transparent allocation keys, creators retain a material stake in the commons. Safadi & Watson [1] show such proportional distribution is feasible, while Floridi & Cowls' [11] principle of responsibility by design grounds why platform providers, not end users, should bear the cost.
- *Pluralistic sourcing and audit respond to epistemic homogenization and the erasure of marginal knowers.* Metrics must track language representation, geographic source distribution, and outlet size/type [17], [51]. Public audits [52] operationalize Medina's [44] call for epistemic virtues of resistance, ensuring diversity is not aspirational but enforceable.

Together, these reforms instantiate a proportionality principle: accessibility and immediacy may defeasibly justify streamlined presentation, but not wholesale suppression of provenance, redistribution, or pluralism. Properly operationalized, they transform epistemic justice from an abstract value into a set of enforceable design and governance obligations [15], [11].

To clarify how these proposals follow directly from the injustices diagnosed above, Table 2 maps the mechanisms identified in the three-layer framework to the specific epistemic harms they generate, the corresponding reforms, and their normative justifications. This mapping shows that the recommendations are not external policy add-ons but principled responses that logically emerge from the justice analysis.

3.5 Proportionality, permissibility, and countervailing goods

Countervailing goods, accessibility for low-literacy users, efficiency in time-critical health contexts, or safety in emergencies are genuine but defeasible [2]. A proportionality analysis clarifies that reduced provenance may be warranted only when immediacy is strictly necessary and only through the least restrictive means: inline citations, expandable provenance, and multi-source answer views [20]. Current zero-click designs fail this test by suppressing provenance more extensively than required, externalizing epistemic costs to upstream creators.

Legal permissibility (e.g., fair use or licensing) cannot settle ethical permissibility [7], [6]. Recognition, accountability, and justice demand governance structures that law

alone does not guarantee. User preference for concise answers likewise cannot justify systematic effacement: preferences shaped under constrained design reflect engineered scarcity, not informed choice. Properly configured, zero-click systems could preserve immediacy while embedding reciprocity and accountability. As currently constituted, governance asymmetries convert accessibility into epistemic enclosure, undermining the pluralism on which a resilient knowledge commons depends.

4 Conclusion: from collapse to reimagining knowledge infrastructures

This article has interrogated how zero-click architectures reconfigure the very conditions of knowledge in the digital age. The guiding question, what it means when platforms no longer direct us outward toward human authorship but instead speak as epistemic agents in their own right, has been addressed through a three-layer framework of infrastructures, appropriation, and governance. What emerges is a portrait of platforms not as neutral intermediaries, but as actors that reorder economies of attention, redistribute epistemic credit upward, and recast the norms of credibility and trust.

The analysis has shown that the apparent convenience of zero-click summaries conceals deeper epistemic harms. Attention is captured, provenance effaced, and the work of creators appropriated in ways that collapse plural contributions into homogenized outputs. Such dynamics instantiate a novel form of technological mediation in which economic disenfranchisement is inseparable from epistemic marginalization. Zero-click collapse, therefore, must be understood not as a mere feature of interface design but as a structural transformation of the knowledge commons with profound philosophical implications.

The stakes are considerable. If left unchallenged, the consolidation of epistemic authority within proprietary infrastructures risks normalizing epistemic injustice, eroding democratic pluralism, and diminishing the conditions of public trust. Yet collapse is not destiny. A human-centered paradigm, exemplified by value-sensitive and participatory governance, gestures toward alternative trajectories: infrastructures that acknowledge provenance, systems that redistribute value, and governance models that embed inclusivity, transparency, and accountability.

The argument advanced here does not deny the legitimate aims of accessibility, speed, or safety; it specifies the conditions under which they defeasibly justify streamlined delivery. A proportionality rule follows: Deploy zero-click summaries with provenance by default, and relax traceability only where (a) there is a demonstrable time-critical or accessibility constraint, (b) inline citations and expandable sources remain visible, (c) plural-source views are available within one click, and (d) any relaxation is logged for audit. This least-restrictive-means principle aligns accessibility with epistemic justice: it preserves the benefits of immediacy without converting exceptional pressures into a standing license to efface authorship.

In closing, this paper has not aimed to provide an exhaustive account of the political economy or legal regulation of generative AI but to delimit a conceptual terrain where philosophy of technology can intervene with clarity and force. By theorizing zero-click collapse as the emergence of platforms as epistemic agents, the analysis contributes to debates on technological mediation, epistemic injustice, and the normative dimensions of design. It shows that attribution, authority, and trust are not ancillary concerns but constitutive features of digital infrastructures. The three-layer framework demonstrates

how, under prevailing design and governance conditions, zero-click AI systematically produces testimonial and hermeneutical injustices. It also shows that where provenance, reciprocity, and pluralistic governance are embedded, the characteristic harms do not arise, underscoring their systemic but contingent nature. The challenge for philosophy, then, is not only to diagnose collapse but also to articulate the ethical horizons within which alternative infrastructures of knowledge can be designed and sustained.

Primarily conceptual and theory-building in orientation, this article forgoes an empirical estimation of zero-click prevalence and a quantitative assessment of the specific harms described. In light of the significant variance in interface design, sourcing practices, and policy regimes, and given the rapid evolution of these technologies, the framework delineates conditional mechanisms rather than advancing universal claims concerning AI-mediated search or summarization. Finally, the reform proposals are articulated as normative thresholds and design obligations rooted in the preceding justice analysis; they are intended as evaluative criteria rather than exhaustive implementation blueprints or formal regulatory impact assessments.

Author contributions
SO is the sole author of this manuscript and was responsible for the conceptualization, theoretical development, analysis, and writing of the manuscript, including the preparation of all figures and tables.

Funding
The author received no specific funding for this work.

Data availability
Data sharing is not applicable to this article, as no datasets were generated or analyzed during the current study.

Code availability
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable. This article does not contain any studies involving human participants or animals performed by the author.

Consent for publication
Not applicable.

Competing interests
The author declares that she has no competing interests.

Received: 3 December 2025 / Accepted: 13 March 2026
Published online: 13 April 2026

References

1. Safadi H, Watson RT. Knowledge monopolies and the innovation divide: a governance perspective. Inf Organ. 2023;33(2):100466. https://doi.org/10.1016/j.infoandorg.2023.100466.
2. Mackenzie A. Digital platforms. In: Felt U, Irwin A, editors. Elgar encyclopedia of science and technology studies. USA: Edward Elgar; 2024. pp. 521–31.
3. Zhang D. Should ChatGPT and Bard share revenue with their data providers? A new business model for the AI era. Adv Artificial Intel Mach Learn. 2023;3(2):1056–93. https://doi.org/10.54364/AAIML.2023.1163.
4. Wu Z, Hu L, & Chen PY. AI and human creativity on content creation: analyzing watermark and detection policies. In: Proceedings of the 58th Hawaii International conference on system sciences, University of Hawai'i at Mānoa, 2025. 4209–18. https://hdl.handle.net/10125/109349
5. Fricker M. Epistemic injustice: power and the ethics of knowing. Oxford: Oxford University Press; 2007.
6. Crawford K. Atlas of AI: power, politics, and the planetary costs of artificial intelligence. USA: Yale University Press; 2021.
7. Pasquale F. *The black box society: the secret algorithms that control money and information*. Harvard University Press, USA; 2015.
8. Bietti E. From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy. In: Proceedings of the 2020 conference on fairness, accountability, and transparency. Association for computing machinery, 2020. 210–9 https://doi.org/10.1145/3351095.3372860
9. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: mapping the debate. Big Data Soc. 2016. https://doi.org/10.1177/2053951716679679.

10. Coeckelbergh M. AI ethics. USA: MIT Press; 2020.
11. Floridi L, Cowls J. A unified framework of five principles for AI in society. In: Floridi L, editor. Ethics, governance, and policies in artificial intelligence, philosophical studies series, vol. 144. Cham: Springer; 2021.
12. Sundar SS. Rise of machine agency: a framework for studying the psychology of human–AI interaction (HAII). J Comput-Mediat Commun. 2020;25(1):74–88. https://doi.org/10.1093/jcmc/zmz026.
13. Verbeek P-P. What things do: philosophical reflections on technology, agency, and design. Pennsylvania: The Pennsylvania State University Press; 2005.
14. Misra S, Barik K. A comprehensive review of human-centric AI, regulatory frameworks, and their role in shaping Industry 5.0. Procedia Computer Sci. 2025;259:1672–81. https://doi.org/10.1016/j.procs.2025.04.122.
15. Verbeek P-P. Moralizing technology: understanding and designing the morality of things. Chicago: The University of Chicago Press; 2011.
16. Ihde D. Technology and the lifeworld: from garden to earth. Indiana: Indiana University Press; 1990.
17. Couldry N, Mejias UA. The costs of connection: how data is colonizing human life and appropriating it for capitalism. Stanford: Stanford University Press; 2019.
18. Bélisle-Pipon J-C, Victor G. Ethics dumping in artificial intelligence. Front Artificial Intell. 2024;7:1426761. https://doi.org/10.3389/frai.2024.1426761.
19. Mourtzis D. Outlook, trends, and future directions toward Industry 5.0. In: Manufacturing from Industry 4.0 to Industry 5.0: advances and applications. 2024. 465–92. https://doi.org/10.1016/B978-0-443-13924-6.00015-6
20. Gao Y, Wang Z, & Huang Y. Pandora Box or Golden Fleece: economic analysis of generative AI adoption on creation platforms. In: Proceedings of the 58th Hawaii international conference on system sciences, University of Hawai'i at Mānoa. 2025. 2678–87. https://hdl.handle.net/10125/109165
21. Berners-Lee T, Fischetti M. Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventor. USA: HarperCollins Publishers; 2000.
22. Ziewitz M, Pentzold C. In search of internet governance: performing order in digitally networked environments. New Media Soc. 2014;16(2):306–22. https://doi.org/10.1177/1461444813504192.
23. Parker G, Van Alstyne M. Innovation, openness & platform control. In: Proceedings of the 11th ACM Conference on Electronic Commerce (EC'10), 2010. 95–6. https://doi.org/10.1145/1807342.1807357
24. Star SL, Ruhleder K. Steps toward an ecology of infrastructure: design and access for large information spaces. Inf Syst Res. 1996;7(1):111–34. https://doi.org/10.1287/isre.7.1.111.
25. Edwards PN. A vast machine: computer models, climate data, and the politics of global warming. USA: The MIT Press; 2013.
26. Hancock JT, Naaman M, Levy K. AI-mediated communication: definition, research agenda, and ethical considerations. J Comput-Mediat Commun. 2020;25(1):89–100. https://doi.org/10.1093/jcmc/zmz022.
27. Gillespie T. Custodians of the internet: platforms, content moderation, and the hidden decisions that shape social media. USA: Yale University Press; 2018.
28. Katsamakas E, Sanchez-Cartas JM. Generative artificial intelligence, content creation, and platforms. J Ind Compet Trade. 2024;24(1):19. https://doi.org/10.1007/s10842-024-00430-9.
29. Thatcher J, O'Sullivan D, Mahmoudi D. Data colonialism through accumulation by dispossession: new metaphors for daily data. Environ Plan D: Soc Space. 2016;34(6):990–1006. https://doi.org/10.1177/0263775816633195.
30. Santos-Gonçalves T. Local media in the age of AI. In: Revolutionizing communication: the role of artificial intelligence, 2024.129–39. https://doi.org/10.1201/9781003473633-12
31. Farhoudinia B, Ozturkcan S, Kasap N. Fake news in business and management literature: a systematic review of definitions, theories, methods, and implications. Aslib J Inf Manag. 2025;77(2):306–29. https://doi.org/10.1108/AJIM-09-2022-0418.
32. Suzor NP. Lawless: the secret rules that govern our digital lives. Cambridge: Cambridge University Press; 2019.
33. Brand DJ. Responsible artificial intelligence in government: development of a legal framework for South Africa. JeDEM - eJournal eDemocracy Open Government. 2022;14(1):130–50. https://doi.org/10.29379/jedem.v14i1.678.
34. Joyce DW, Kormilitzin A, Hamer-Hunt J, McKee KR, Tomasev N. Defining acceptable data collection and reuse standards for queer artificial intelligence research in mental health: protocol for the online PARQAIR-MH Delphi study. BMJ Open. 2024;14(3):e079105. https://doi.org/10.1136/bmjopen-2023-079105.
35. Feenberg A. Transforming technology: a critical theory revisited. Oxford: Oxford University Press; 2002.
36. Jasanoff S, editor. States of knowledge: the co-production of science and the social order. UK: Routledge; 2006.
37. Miernicki M, Ng I. Machines, attribution, and integrity: artificial intelligence and moral rights. Jusletter IT; 2019.
38. Mazzi F. Authorship in artificial intelligence-generated works: Exploring originality in text prompts and artificial intelligence outputs through philosophical foundations of copyright and collage protection. J World Intellect Prop. 2024;27(3):410–27. https://doi.org/10.1111/jwip.12310.
39. Kalyvaki M, Nash K, McIntosh H. AI, copyright, and business: navigating global legal challenges in the era of generative content and digital replicas. Int Rev Law, Computers Technol. 2025. https://doi.org/10.1080/13600869.2025.2590795.
40. Chamma D, Ahmadpour N, Ishtiaque Ahmed S, Jahan Mim N, & Zhang WQ. Epistemic injustice in and through AI. In: OZCHI 2025: Proceedings of the 37th Australian conference on human-computer interaction. 2025. https://doi.org/10.1145/3764687.3767279
41. Faissner M, Lenk J, Müller R. AI and epistemic injustice. In: Hähnel M, Müller R, editors. A companion to applied philosophy of AI. USA: John Wiley & Sons; 2025.
42. Dotson K. A cautionary tale: on limiting epistemic oppression. Front: A J Women Stud. 2012;33(1):24–47. https://doi.org/10.5250/fronjwomestud.33.1.0024.
43. Pohlhaus G. Relational knowing and epistemic injustice: toward a theory of willful hermeneutical ignorance. Hypatia. 2012;27(4):715–35. https://doi.org/10.1111/j.1527-2001.2011.01222.x.
44. Medina J. The epistemology of resistance: gender and racial oppression, epistemic injustice, and the social imagination. Oxford: Oxford University Press; 2013.
45. Alcoff LM. Epistemologies of ignorance: three types. In: Sullivan S, Tuana N, editors. Race and epistemologies of ignorance. NY: SUNY Press; 2007. pp. 39–58.
46. Mills CW. White ignorance. In: Sullivan S, Tuana N, editors. Race and epistemologies of ignorance. New York: SUNY Press; 2007. pp. 11–38.

47. Ozturkcan S. (forthcoming) Choice without choosing: algorithmic marketing and the illusion of consumer autonomy. In: Pardalos PM, Skandali D & Tsioufis M, editors. Springer handbook on AI and optimization in marketing. Springer: Berlin
48. Alcoff L. Real knowing: new versions of the coherence theory. USA: Cornell University Press; 1996.
49. Netipatalachoochote S, Pailler L. Developing artificial intelligence legislation in thailand: lessons from the European Union. J Human Rights, Culture Legal Syst. 2025;5(1):424. https://doi.org/10.53955/jhcls.v5i1.424.
50. Ozturkcan S, Bozdağ AA. Responsible AI in marketing: AI booing and AI washing cycle of AI mistrust. Int J Mark Res. 2025;67(6):696–722. https://doi.org/10.1177/14707853251379285.
51. Helberger N, Wojcieszak M. Exposure diversity. In: Napoli PM, editor. Mediated communication. Berlin: De Gruyter Mouton; 2018. pp. 535–60.
52. Raftopoulos, M. Designing tiny robots: engaging stakeholders in collaborative AI system design. In: ACM International conference proceeding series, 2024. 121–32. https://doi.org/10.1145/3681716.3681722

Publisher's Note